

MEMBRANE TRANSPORT PROTEINS**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of membrane transport proteins and to the use of these sequences in the diagnosis, treatment, and prevention of membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer.

BACKGROUND OF THE INVENTION

Eukaryotic cells are bound by a lipid bilayer membrane and subdivided into functionally distinct, membrane bound compartments. The membranes maintain essential differences between the cytosol, the extracellular environment, and the contents of intracellular organelles such as the Golgi or the endoplasmic reticulum. As lipid membranes are highly impermeable to most polar molecules, transport of essential nutrients; metal ions such as K⁺, NH₄⁺, P_i, SO₄²⁻; sugars; vitamins; metabolic waste products; cell signaling molecules; drugs; peptides; and proteins and other macromolecules across lipid membranes and between organelles must be mediated by a variety of transport molecules. Many transport mechanisms are substrate specific, with each transport protein carrying particular members of a molecular class, such as ions, sugars, or amino acids, across membranes. For example, amino acids are imported into cells via specific amino acid permeases.

Transport proteins are multi-pass transmembrane proteins, which either actively transport molecules across the membrane or passively allow them to cross. Active transport involves directional pumping of a solute across the membrane, usually against an electrochemical gradient. Active transport is tightly coupled to a source of metabolic energy, such as ATP hydrolysis or an electrochemically favorable ion gradient. Passive transport involves the movement of a solute down its electrochemical gradient. Transport proteins can be further classified as either carrier proteins or channel proteins. Carrier proteins, which can function in active or passive transport, bind to a specific solute to be transported and undergo a conformational change which transfers the bound solute across the membrane. Channel proteins, which only function in passive transport, form hydrophilic pores across the membrane. When the pores open, specific solutes, such as inorganic ions, pass through the membrane and down the electrochemical gradient of the solute.

Transport proteins play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). A variety of human inherited diseases are caused by mutation of transport proteins. For example, cystinuria is an inherited disease that results from the inability to transport cystine, the

disulfide-linked dimer of cysteine, from the urine into the blood. Accumulation of cystine in the urine leads to the formation of cystine stones in the kidneys. Also, many transport proteins are composed of subunits that may confer specificity for the tissue in which the transport mechanism functions, and are therefore associated with tissue-specific disorders. Examples of transport proteins include
5 facilitative transporters, the secondary active symporters and antiporters driven by ion gradients, and active ATP binding cassette transporters involved in multiple-drug resistance and targeting of antigenic peptides to MHC Class I molecules, and the E1-E2 cation transport ATPases.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with
10 simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems wherein the movement of sodium into the cell down its electrochemical gradient co-transport a second solute into the cell. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous Na⁺/K⁺ ATPase. Sodium-coupled transporters include
15 the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). These three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging techniques and for specific targeting of radioisotopes to the thyroid
20 gland (Levy, O. et al. (1997) Proc. Natl. Acad. Sci. USA 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) J. Biol. Chem. 273:7501-7506).

The largest and most diverse family of transport proteins is the ATP-binding cassette (ABC) transporters. As a family, ABC transporters can transport substances that differ markedly in chemical
25 structure and size, ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. Each ABC transporter consists of four modules: two nucleotide-binding domains (NBDs), which hydrolyze ATP to supply the energy required for transport; and two membrane-spanning domains (MSDs), which may form membrane channels. The NBDs consist of approximately two hundred conserved amino
30 acid residues while the MSDs each contain six putative transmembrane segments. (See, e.g., Saurin, W. et al. (1994) Mol. Microbiol. 12:993-1004; Shani, N. et al. (1996) J. Biol. Chem. 271:8725-8730; Koster, W. and B. Bohm (1992) Mol. Gen. Genet. 232:399-407.) The four ABC transporter modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane conductance regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product

contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system associated with antigen processing (Androlewicz, M.J. et al. (1994) Proc. Natl. Acad. Sci. USA 91:12716-12720).

5 Several genetic diseases are attributed to defects in ABC transporters, including the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel; Welsh, M.J. and A.E. Smith (1993) Cell 73:1251-1254); X-linked adrenoleukodystrophy, an inborn error of peroxisomal β -oxidation of very long chain fatty acids (adrenoleukodystrophy protein, ALDP); Zellweger syndrome, an inborn error of peroxisome biogenesis (peroxisomal membrane protein-70, PMP70); and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). The ABC transporters known as P-glycoproteins, or multidrug resistance (MDR) proteins, are associated with resistance to a wide range of hydrophobic drugs (MDR1; Gottesman, M.M. and I. Pastan (1993) Annu. Rev. Biochem. 62:385-427) or with phosphatidylcholine transport (MDR2; Ruetz, S. and P. Gros (1994) Cell 77:1071-1081). MDR is common in cancer cells, and contributes to low efficacy or failure of chemotherapy (Taglight, D. and S. Michaelis (1998) Methods Enzymol. 292:131-163). MDR is mediated by transporters, e.g., P-glycoproteins or the multidrug resistance-associated protein MRP, that normally function in the liver, intestines, and kidney to move toxic substances from the cytosol into the bile, intestinal lumen, or urine. In cancerous cells, these transporters extrude chemotherapeutic agents into the extracellular space, thereby conferring drug resistance. Recently, an ABC transporter-type protein was isolated from a human leukemia cell line. This transporter, termed the anthracycline resistance associated protein (GI 1279457, SEQ ID NO:42), is overexpressed in a multidrug resistant leukemia cell sub-line, and has sequence homology with other multidrug-resistance associated proteins including MRP (Longhurst, T.J. et al. (1996) Br. J. Cancer 74:1331-1335).

25 Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant fraction of fatty acid transport appears to occur via a high affinity, low capacity protein-mediated transport process. Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as 30 muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates the cells' uptake of long-chain fatty acids. Expression studies suggest a role for FATP in lipid metabolism, obesity, and type II diabetes mellitus (Hui, T.Y. et al. (1998) J. Biol. Chem. 273:27420-27429).

E1-E2 (or P-type) ATPases constitute a superfamily of cation transporters present in both

prokaryotes and eukaryotes that mediate membrane flux of all biologically relevant cations. These ATPases are postulated to exist in two different conformational states, designated E1 and E2, during the course of the ATP hydrolysis reaction, and to conserve the energy from ATP hydrolysis in the form of an acyl phosphate, primarily an aspartyl phosphate. Members of this family are divided into 5 four major groups; the Ca^{+2} -transporting ATPases, Na^+/K^+ -and gastric H^+/K^+ -transporting ATPases, plasma membrane H^+ -transporting ATPases (proton pumps), and the bacterial P-type ATPases (BLOCKS: BL00154, P-type cation-transporting ATPase superfamily signature).

The metabolism of amino acids is complex and highly regulated. While cells are capable of creating most amino acids de novo, the import of amino acids into cells via specific amino acid 10 permease proteins is vital for maintaining the appropriate and complete availability of all necessary amino acids. This is particularly important during cell proliferation and differentiation. In addition to their role as protein building blocks, amino acids also serve as precursors for a variety of other important macromolecules. For example, the hormone thyroxine, the pigment melanin, and the neurotransmitters histamine, epinephrine, and serotonin are produced from various amino acid 15 precursors, including histidine, tyrosine, and tryptophan. A component of sphingolipid formation, sphingosine, is derived from serine. Porphyrin rings, which are components of heme molecules, use glycine as a nitrogen donor. Significant portions of the ring structures of purines and pyrimidines, components of nucleic acids, are formed from the breakdown of numerous amino acids. Amino acids are also important in energy metabolism. Unlike fatty acids and glucose, amino acids cannot be 20 stored in the cell, so excess amino acids are fed into the citric acid cycle to produce energy molecules including fatty acids, ketone bodies, and glucose. Thus, precise control of amino acid metabolism is extremely important to both proliferating and non-proliferating cells.

The E16 gene, cloned from human peripheral blood lymphocytes, encodes a 241 amino acid integral membrane protein with multiple predicted transmembrane domains (Gaugitsch, H.W. et al. 25 (1992) J. Biol. Chem. 267:11267-11273). E16 gene expression is closely linked to cellular activation and division. In myeloid and lymphoid cells, E16 transcripts are rapidly induced and rapidly degraded after stimulation. This pattern of expression resembles the kinetics seen for proto-oncogenes and lymphokines in the T cell system. Elevated levels of E16 expression were detected in colonic, gastric, and breast adenocarcinomas, and in lymphoma, while little or no E16 expression was 30 detected in normal (non-cancerous) human tissues such as adult brain, lung, liver, colon, esophagus, stomach, or kidney, nor in four-month fetal brain, lung, liver, or kidney (Wolf, D.A. et al. (1996) Cancer Res. 56:5012-5022; Gaugitsch et al., *supra*). E16 was detected in every cell line tested. Its presence in rapidly dividing cell lines and its absence in human tissues with low proliferative potential suggest that E16 is directly involved in the cell division process, where it helps provide important

building blocks for energy metabolism, biochemical synthetic pathways, and protein synthesis.

Post-translational modification of polypeptides occurs in the lumen of the Golgi apparatus. Such modifications include, for example, the addition of sugar molecules by enzymes such as N-acetylglucosaminyltransferase, to produce glycoproteins. The sugar-donating molecules in this reaction are typically nucleotide sugars, such as uridine diphosphate-galactose (UDP-Gal). UDP-Gal and other nucleotide sugars are transported from the cytosol into the Golgi apparatus by specific transporter molecules. The availability of these nucleotide sugars can regulate which glycoproteins are synthesized, and therefore has a significant impact on cellular function (Toma, L. et al. (1996) J. Biol. Chem. 271:3897-3901; Guillen, E. et al. (1998) Proc. Natl. Acad. Sci. USA 95:7888-7892).

The discovery of new membrane transport proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer.

15

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, membrane transport proteins, referred to collectively as "MTRP" and individually as "MTRP-1," "MTRP-2," "MTRP-3," "MTRP-4," "MTRP-5," "MTRP-6," "MTRP-7," "MTRP-8," "MTRP-9," "MTRP-10," "MTRP-11," "MTRP-12," "MTRP-13," "MTRP-14," "MTRP-15," "MTRP-16," and "MTRP-17." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also includes a polypeptide comprising an amino acid sequence that differs by one or more conservative amino acid substitutions from an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The

invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of: (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing a polynucleotide of the invention under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-17 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of MTRP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially

purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of MTRP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figures 1A, 1B, 1C, and 1D show the amino acid sequence alignment between MTRP-3 (Incyte Clone ID 1720440; SEQ ID NO:3) and mouse fatty acid transport protein (GI 2612939; SEQ ID NO:35), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Figures 2A, 2B, 2C, and 2D show the amino acid sequence alignment between MTRP-4 (Incyte Clone ID 2274290; SEQ ID NO:4) and Schistosoma mansoni SMDR1 (GI 425474; SEQ ID NO:36), produced using the multisequence alignment program of LASERGENE software (DNASTAR).

Figures 3A, 3B, 3C, and 3D show the amino acid sequence alignment between MTRP-5 (Incyte Clone ID 2740029; SEQ ID NO:5) and rat sodium-dependent multivitamin transporter (GI 3015617; SEQ ID NO:37), produced using the multisequence alignment program of LASERGENE software (DNASTAR).

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOS), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding MTRP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of MTRP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding MTRP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze MTRP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will 5 be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so 10 forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now 15 described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

20 "MTRP" refers to the amino acid sequences of substantially purified MTRP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of 25 MTRP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MTRP either by directly interacting with MTRP or by acting on components of the biological pathway in which MTRP participates.

An "allelic variant" is an alternative form of the gene encoding MTRP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in 30 polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding MTRP include those sequences with deletions,

insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MTRP or a polypeptide with at least one functional characteristic of MTRP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding MTRP, and improper or unexpected hybridization to allelic variants, 5 with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding MTRP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent MTRP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the 10 residues, as long as the biological or immunological activity of MTRP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, 15 isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino 20 acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

25 The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of MTRP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of MTRP either by directly interacting with MTRP or by acting on components of the biological pathway in which MTRP participates.

30 The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind MTRP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the

translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

5 The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen
10 used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes
15 and to block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand.

20 The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic MTRP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

25 The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization
30 between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acid strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding MTRP or fragments of MTRP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be

associated with a stabilizing agent such as a carbohydrate.. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

“Consensus sequence” refers to a nucleic acid sequence which has been resequenced to 5 resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of one or more Incyte Clones and, in some cases, one or more public domain ESTs, using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

10 “Conservative amino acid substitutions” are those substitutions that, when made, least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
15	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
20	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
25	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
30	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
35	Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

40 A “deletion” refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "fragment" is a unique portion of MTRP or the polynucleotide encoding MTRP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:18-34 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:18-34, for example, as distinct from any other sequence in the same genome. A fragment of SEQ ID NO:18-34 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:18-34 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:18-34 and the region of SEQ ID NO:18-34 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-17 is encoded by a fragment of SEQ ID NO:18-34. A fragment of SEQ ID NO:1-17 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-17. For example, a fragment of SEQ ID NO:1-17 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-17. The precise length of a fragment of SEQ ID NO:1-17 and the region of SEQ ID NO:1-17 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially

complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequence pairs.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2

Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such default parameters may be, for example:

5 *Matrix: BLOSUM62*
10 *Reward for match: 1*
 Penalty for mismatch: -2
 Open Gap: 5 and Extension Gap: 2 penalties
 Gap x drop-off: 50
 Expect: 10
 Word Size: 11
 Filter: on

15 Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported
20 by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

25 Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

30 The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the hydrophobicity and acidity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of

polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and “diagonals saved”=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the “percent similarity” between aligned polypeptide sequence pairs.

5 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the “BLAST 2 Sequences” tool Version 2.0.9 (May-07-1999) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

10 *Gap x drop-off: 50*

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence,
15 for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be
20 used to describe a length over which percentage identity may be measured.

“Human artificial chromosomes” (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term “humanized antibody” refers to antibody molecules in which the amino acid
25 sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

“Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of identity. Specific
30 hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and

may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml denatured salmon sperm DNA.

5 Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Generally, such wash temperatures are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and
10 conditions for nucleic acid hybridization are well known and can be found in Sambrook et al., 1989,
Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY;
specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour.

15 Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for
20 RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid
25 sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., $C_{0,t}$ or $R_{0,t}$ analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

30 The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect

cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

5 The term "modulate" refers to a change in the activity of MTRP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of MTRP.

10 The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

15 "Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

20 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

25 "Probe" refers to nucleic acid sequences encoding MTRP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule.

Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

30 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the

specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques

such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to 5 transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic 10 acids encoding MTRP, or fragments thereof, or MTRP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or 15 synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

20 The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by 25 different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

30 "Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection,

electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

5 A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or
10 greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the
15 reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The
20 presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human membrane transport proteins (MTRP),
30 the polynucleotides encoding MTRP, and the use of these compositions for the diagnosis, treatment, or prevention of membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding MTRP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide

and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each MTRP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each MTRP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

As shown in Figures 1A, 1B, 1C, and 1D, MTRP-3 has chemical and structural similarity with mouse fatty acid transport protein (FATP; GI 2612939; SEQ ID NO:35). In particular, MTRP-3 and FATP share 65% identity. As shown in Figures 2A, 2B, 2C, and 2D, MTRP-4 has chemical and structural similarity with *Schistosoma mansoni* ATP-binding cassette family protein, SMDR-1 (GI 425474; SEQ ID NO:36). In particular, MTRP-4 and SMDR-1 share 38% identity. As shown in Figures 3A, 3B, 3C, and 3D, MTRP-5 has chemical and structural similarity with rat sodium-dependent multivitamin transporter (SMVT; GI 3015617; SEQ ID NO:37). In particular, MTRP-5 and SMVT share 82% identity.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding MTRP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists tissue categories which express MTRP as a fraction of total tissues expressing MTRP. Column 3 lists diseases, disorders, or conditions associated with those tissues expressing MTRP as a fraction of total tissues expressing MTRP. Column 4 lists the vectors used to subclone each cDNA library.

Of particular note are the expression patterns of SEQ ID NO:30 and SEQ ID NO:31. SEQ ID NO:30 is expressed in only five libraries, of which at least four (80%) are associated with cell proliferation and at least one (20%) with inflammation. Two (40%) of the five libraries are associated with cardiovascular tissue, and one each (20%) with gastrointestinal, nervous, and reproductive tissues. SEQ ID NO:31 is expressed in only four libraries, of which at least three (75%) are associated with cell proliferation and at least two (50%) with inflammation or the immune response.

Two (50%) of the four libraries are associated with hematopoietic/immune tissue, and one each (25%) with cardiovascular and reproductive tissues.

The following fragments of the nucleotide sequences encoding MTRP are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:18-34 and to distinguish between SEQ ID NO:18-34 and related polynucleotide sequences. The useful fragments include the fragment of SEQ ID NO:18 from about nucleotide 110 to about nucleotide 154; the fragment of SEQ ID NO:19 from about nucleotide 759 to about nucleotide 839; the fragment of SEQ ID NO:20 from about nucleotide 1531 to about nucleotide 1578; the fragment of SEQ ID NO:21 from about nucleotide 538 to about nucleotide 597; the fragment of SEQ ID NO:22 from about nucleotide 2241 to about nucleotide 2294; the fragment of SEQ ID NO:23 from about nucleotide 116 to about nucleotide 145; the fragment of SEQ ID NO:24 from about nucleotide 60 to about nucleotide 89; the fragment of SEQ ID NO:25 from about nucleotide 160 to about nucleotide 189; the fragment of SEQ ID NO:26 from about nucleotide 763 to about nucleotide 792; the fragment of SEQ ID NO:27 from about nucleotide 43 to about nucleotide 72; the fragment of SEQ ID NO:28 from about nucleotide 361 to about nucleotide 405; the fragment of SEQ ID NO:29 from about nucleotide 35 to about nucleotide 79; the fragment of SEQ ID NO:30 from about nucleotide 206 to about nucleotide 250; the fragment of SEQ ID NO:31 from about nucleotide 71 to about nucleotide 115; the fragment of SEQ ID NO:32 from about nucleotide 161 to about nucleotide 205; the fragment of SEQ ID NO:33 from about nucleotide 364 to about nucleotide 408; and the fragment of SEQ ID NO:34 from about nucleotide 18 to about nucleotide 62. The polypeptides encoded by the specified fragments of SEQ ID NO:20-30 and SEQ ID NO:32-34 are useful, for example, as immunogenic peptides.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding MTRP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses MTRP variants. A preferred MTRP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the MTRP amino acid sequence, and which contains at least one functional or structural characteristic of MTRP.

The invention also encompasses polynucleotides which encode MTRP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:18-34, which encodes MTRP.

The invention also encompasses a variant of a polynucleotide sequence encoding MTRP. In particular, such a variant polynucleotide sequence will have at least about 75%, or alternatively at

least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding MTRP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:18-34 which has at least about 75%, or alternatively at least about 85%, or even at least about 5 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:18-34. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of MTRP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding MTRP, some bearing minimal 10 similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These 15 combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring MTRP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode MTRP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring MTRP under appropriately selected 20 conditions of stringency, it may be advantageous to produce nucleotide sequences encoding MTRP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide 25 occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding MTRP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode MTRP and MTRP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce 30 mutations into a sequence encoding MTRP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:18-34 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol.

152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment 5 of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), 10 PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short 15 Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding MTRP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, 20 restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids 25 Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences 30 are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National

Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been
5 size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze
10 the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process
15 from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode MTRP may be cloned in recombinant DNA molecules that direct expression of MTRP,
20 or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express MTRP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter MTRP-encoding sequences for a variety of purposes including, but
25 not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

30 In another embodiment, sequences encoding MTRP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, MTRP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g.,

Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of MTRP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

5 The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

10 In order to express a biologically active MTRP, the nucleotide sequences encoding MTRP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences 15 encoding MTRP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding MTRP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding MTRP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals 20 may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994)

25 Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding MTRP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A 30 Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding MTRP. These include, but are not limited to, microorganisms such as bacteria transformed

with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or 5 animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding MTRP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding MTRP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSORT1 10 plasmid (Life Technologies). Ligation of sequences encoding MTRP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. 15 Chem.* 264:5503-5509.) When large quantities of MTRP are needed, e.g. for the production of antibodies, vectors which direct high level expression of MTRP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of MTRP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH 20 promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

25 Plant systems may also be used for expression of MTRP. Transcription of sequences encoding MTRP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock 30 promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases

where an adenovirus is used as an expression vector, sequences encoding MTRP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses MTRP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. 5 Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are 10 constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of MTRP in cell lines is preferred. For example, sequences encoding MTRP can be transformed into 15 cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express 20 the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et 25 al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) 30 J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to

quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding MTRP is inserted within a marker gene sequence, transformed cells containing sequences encoding MTRP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding MTRP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding MTRP and that express MTRP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of MTRP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on MTRP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding MTRP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding MTRP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for

ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding MTRP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode MTRP may be designed to contain signal sequences which direct secretion of MTRP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "pro" or "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding MTRP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric MTRP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MTRP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the MTRP encoding sequence and the heterologous protein sequence, so that MTRP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of

fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled MTRP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the 5 T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

Fragments of MTRP may be produced not only by recombinant means, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be 10 achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of MTRP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of MTRP and membrane transport proteins, including amino acid transporters, ABC 15 transporters, nucleotide-sugar transporters, transmembrane carrier proteins, and ATP-dependent transporter proteins. In addition, the expression of MTRP is closely associated with nervous, reproductive, and gastrointestinal tissues; cancer and other cell proliferative conditions; and with inflammation and the immune response. Therefore, MTRP appears to play a role in membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer. 20 In the treatment of disorders associated with increased MTRP expression or activity, it is desirable to decrease the expression or activity of MTRP. In the treatment of disorders associated with decreased MTRP expression or activity, it is desirable to increase the expression or activity of MTRP.

Therefore, in one embodiment, MTRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or 25 activity of MTRP. Examples of such disorders include, but are not limited to, a membrane transport disorder such as cystinuria, dibasicaminoaciduria, hypercystinuria, lysinuria, hartnup disease, tryptophan malabsorption, methionine malabsorption, histidinuria, iminoglycinuria, dicarboxylicaminoaciduria, cystinosis, renal glycosuria, glucose-galactose malabsorption, familial hypercholesterolemia, hypouricemia, familial hypophosphatemic rickets, congenital chloridorrhea, 30 cystic fibrosis, familial goiter, distal renal tubular acidosis, Menkes' disease, lethal diarrhea, nephrogenic diabetes insipidus, juvenile pernicious anemia, folate malabsorption, adrenoleukodystrophy, hereditary myoglobinuria, Zellweger syndrome, hyperinsulinemic hypoglycemia, akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus,

diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, and prostate cancer; a cardiac disorder associated with transport such as angina,

5 bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, and polymyositis; a neurological disorder associated with transport such as Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and

10 schizophrenia; and an other disorder associated with transport such as neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, and hypercholesterolemia; an immune/inflammatory disorder such as

15 acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia

20 with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic

25 lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation; a viral, bacterial, fungal, parasitic, protozoal, or helminthic infection; and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis,

30 primary thrombocythemia; and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma; and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing MTRP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP including, but not limited to, those described above.

- In a further embodiment, a pharmaceutical composition comprising a substantially purified MTRP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP including, but not limited to, those provided above.

- In still another embodiment, an agonist which modulates the activity of MTRP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MTRP including, but not limited to, those listed above.

- In a further embodiment, an antagonist of MTRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MTRP. Examples of such disorders include, but are not limited to, those membrane transport disorders; immune/inflammatory disorders; and cell proliferative disorders including cancer described above. In one aspect, an antibody which specifically binds MTRP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express MTRP.

- In an additional embodiment, a vector expressing the complement of the polynucleotide encoding MTRP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MTRP including, but not limited to, those described above.

- In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

- An antagonist of MTRP may be produced using methods which are generally known in the art. In particular, purified MTRP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind MTRP. Antibodies to MTRP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with MTRP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, 5 BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to MTRP have an amino acid sequence consisting of at least about 5 amino acids, and generally will 10 consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of MTRP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

15 Monoclonal antibodies to MTRP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and 20 Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, 25 S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce MTRP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

30 Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for MTRP may also be generated.

For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) *Science* 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between MTRP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MTRP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MTRP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of MTRP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple MTRP epitopes, represents the average affinity, or avidity, of the antibodies for MTRP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular MTRP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10⁹ to 10¹² L/mole are preferred for use in immunoassays in which the MTRP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10⁶ to 10⁷ L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of MTRP, preferably in active form, from the antibody (Catty, D. (1988) *Antibodies, Volume I: A Practical Approach*, IRL Press, Washington, DC; Liddell, J.E. and Cryer, A. (1991) *A Practical Guide to Monoclonal Antibodies*, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of MTRP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding MTRP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding MTRP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding MTRP. Thus, complementary molecules or fragments may be used to modulate MTRP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding MTRP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding MTRP. (See, e.g., Sambrook, *supra*; Ausubel, 1995, *supra*.)

Genes encoding MTRP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding MTRP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding MTRP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may be employed. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme

molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding MTRP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding MTRP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterate linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of MTRP, antibodies to MTRP, and mimetics, agonists, antagonists, or inhibitors of MTRP. The compositions 5 may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any 10 number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable 15 pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using 20 pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active 25 compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, 30 such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to

characterize the quantity of active compound, i.e., dosage..

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of MTRP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example MTRP or fragments thereof, antibodies of MTRP, and agonists, antagonists or inhibitors of MTRP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind MTRP may be used for the diagnosis of disorders characterized by expression of MTRP, or in assays to monitor patients being

treated with MTRP or agonists, antagonists, or inhibitors of MTRP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for MTRP include methods which utilize the antibody and a label to detect MTRP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and 5 may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring MTRP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of MTRP expression. Normal or standard values for MTRP expression are established by combining body fluids or cell extracts 10 taken from normal mammalian subjects, for example, human subjects, with antibody to MTRP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of MTRP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

15 In another embodiment of the invention, the polynucleotides encoding MTRP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of MTRP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess 20 expression of MTRP, and to monitor regulation of MTRP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding MTRP or closely related molecules may be used 25 to identify nucleic acid sequences which encode MTRP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding MTRP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the MTRP encoding sequences. The hybridization probes of the subject 30 invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:18-34 or from genomic sequences including promoters, enhancers, and introns of the MTRP gene.

Means for producing specific hybridization probes for DNAs encoding MTRP include the cloning of polynucleotide sequences encoding MTRP or MTRP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may

be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

5 Polynucleotide sequences encoding MTRP may be used for the diagnosis of disorders associated with expression of MTRP. Examples of such disorders include, but are not limited to, a membrane transport disorder such as cystinuria, dibasicaminoaciduria, hypercystinuria, lysinuria, hartnup disease, tryptophan malabsorption, methionine malabsorption, histidinuria, iminoglycinuria, dicarboxylicaminoaciduria, cystinosis, renal glycosuria, glucose-galactose malabsorption, familial 10 hypercholesterolemia, hypouricemia, familial hypophosphatemic rickets, congenital chloridorrhea, cystic fibrosis, familial goiter, distal renal tubular acidosis, Menkes' disease, lethal diarrhea, nephrogenic diabetes insipidus, juvenile pernicious anemia, folate malabsorption, adrenoleukodystrophy, hereditary myoglobinuria, Zellweger syndrome, hyperinsulinemic 15 hypoglycemia, akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, and prostate cancer; a cardiac disorder associated with transport such as angina, 20 bradycardia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, and polymyositis; a neurological disorder associated with transport such as Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and 25 schizophrenia; and an other disorder associated with transport such as neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, and hypercholesterolemia; an immune/inflammatory disorder such as 30 acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia

with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation; a viral, bacterial, fungal, parasitic, protozoal, or helminthic infection; and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; and a cancer including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma; and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding MTRP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered MTRP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding MTRP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding MTRP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding MTRP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of MTRP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding MTRP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from

normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

5 Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

10 With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development
15 or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding MTRP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding MTRP, or a fragment of a polynucleotide complementary to the polynucleotide encoding
20 MTRP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of MTRP include radiolabeling or biotinylation nucleotides, coamplification of a control nucleic acid, and interpolating results from
25 standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

30 In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and

monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. 5 (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding MTRP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to 10 artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome 15 mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding MTRP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. 20 The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, 25 may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping 30 to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, MTRP, its catalytic or immunogenic fragments, or

oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between MTRP and the agent being tested may be measured.

- 5 Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with MTRP, or fragments thereof, and washed. Bound MTRP is then detected by methods well known in the art. Purified MTRP can
10 also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding MTRP specifically compete with a test compound for binding MTRP.

- 15 In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with MTRP.

- In additional embodiments, the nucleotide sequences which encode MTRP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such
20 properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder
25 of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0633 P, filed November 4, 1998], U.S. Ser. No. [Attorney Docket No. PF-0645 P, filed November 24, 1998], U.S. Ser. No. [Attorney Docket No. PF-0657 P, filed December 22, 1998], and U.S. Ser. No. 60/121,896, are hereby expressly incorporated
30 by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some

tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either 5 isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was 10 isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the 15 recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300- 1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs 20 were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or 25 ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge 30 Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a

high-throughput format (Rao, V.B. (1994) *Anal. Biochem.* 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSCAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA

sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire
5 annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying
10 against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and
15 amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:18-34. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a
20 gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel,
1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This
25 analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

30 The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding MTRP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, 5 reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

10 **V. Extension of MTRP Encoding Polynucleotides**

The full length nucleic acid sequences of SEQ ID NO:18-34 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using 15 OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one 20 extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme 25 (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; 30 Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II

(Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

5 The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones
10 were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

15 The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

20 In like manner, the nucleotide sequences of SEQ ID NO:18-34 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

25 Hybridization probes derived from SEQ ID NO:18-34 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a

SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

5 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and
10 compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, *supra*.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, 15 UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned 20 images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or 25 selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by 30 procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the MTRP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring MTRP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same

procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of MTRP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the MTRP-encoding transcript.

IX. Expression of MTRP

Expression and purification of MTRP is achieved using bacterial or virus-based expression systems. For expression of MTRP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MTRP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MTRP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding MTRP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, MTRP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from MTRP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified MTRP obtained by these methods can be used directly in the

following activity assay.

X. Demonstration of MTRP Activity

ATPase activity associated with MTRP can be measured by hydrolysis of radiolabeled ATP-[γ -³²P], separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered ³²P using a scintillation counter. The reaction mixture contains ATP-[γ -³²P] and varying amounts of MTRP in a suitable buffer incubated at 37°C for a suitable period of time. The reaction is terminated by acid precipitation with trichloroacetic acid and then neutralized with base, and an aliquot of the reaction mixture is subjected to membrane or filter paper-based chromatography to separate the reaction products. The amount of ³²P liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of MTRP in the assay.

MTRP transport activity is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with MTRP mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 1mM Na₂HPO₄, 5 mM Hepes, 3.8 mM NaOH , 50 μ g/ml gentamycin, pH 7.8) to allow expression of MTRP protein. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 10 mM Hepes/Tris, pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, and neurotransmitters) is initiated by adding a ³H-labeled substrate to the oocytes. After 30 minutes of incubation, uptake is terminated by washing the oocytes three times in Na⁺-free medium. Incorporation of ³H is measured, and compared with controls. MTRP transport activity is proportional to the level of internalized ³H-labeled substrate.

XI. Functional Assays

MTRP function is assessed by expressing the sequences encoding MTRP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of

fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; 5 alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of MTRP on gene expression can be assessed using highly purified 10 populations of cells transfected with sequences encoding MTRP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the 15 art. Expression of mRNA encoding MTRP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of MTRP Specific Antibodies

MTRP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., 20 Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the MTRP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well 25 described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-MTRP activity by, for example, binding the peptide or MTRP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG. 30

XIII. Purification of Naturally Occurring MTRP Using Specific Antibodies

Naturally occurring or recombinant MTRP is substantially purified by immunoaffinity

chromatography using antibodies specific for MTRP. An immunoaffinity column is constructed by covalently coupling anti-MTRP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

- 5 Media containing MTRP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MTRP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MTRP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MTRP is collected.

10 **XIV. Identification of Molecules Which Interact with MTRP**

- MTRP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MTRP, washed, and any wells with labeled MTRP complex are assayed. Data obtained using different concentrations 15 of MTRP are used to calculate values for the number, affinity, and association of MTRP with the candidate molecules.

- Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention.
- 20 Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	18	961344	BRSTTUT03	642329R1 (BRSTNOT03), 961344H1 (BRSTTUT03), 3149970H1 (ADRENON04), 3255621H1 (OVARTUN01), 4265773H1 (KIDNNOT32), 4641320H1 (PROSTMTO3), 4875558H1 (COLDNOT01)
2	19	3128782	LUNGUT12	148511X20R1 (FIBRNGT01), 580891H1 (BRAVXT05), 1304328F1 (PLACNOT02), 1469890T6 (PANCTUT02), 1799816T6 (COLNNOT27), 1985910R6 (LUNGAST01), 2722244F6 (LUNGUT10), 3128782F6 (LUNGUT12), 3128782H1 (LUNGUT12), 3276993F6 (PROSBPT06), SBMA03256F1
3	20	1720440	BLADNOT06	744485R1 (BRAITUT01), 859118R1 (BRAITUT03), 944049T1 (ADRENOT03), 1432755R1 (BEPINNON01), 1720440H1 (BLADNOT06), 1808737X11C1 (PROSTUT12), 1812106F6 (PROSTUT12), 2192988E6 (THYRTUT03), 2192988X13F1 (THYRTUT03), 2192988X14F1 (THYRTUT03), 3384757H1 (ESOGNOT04)
4	21	2274290	PROSNON01	1732422F6 (BRSTTUT08), 2098563H1 (BRAITUT02), 2274290H1 (PROSNON01), 2274290X326D2 (PROSNON01), 2598580F6 (UTRSNOT10), 2779864F6 (OVARTUT03), 2864759F6 (KIDNNOT20), 2864759T6 (KIDNNOT20), 3221871H1 (COLNNON03)
5	22	2740029	BRSTTUT14	966363H1 (BRSTNOT05), 1000112R1 (BRSTNOT03), 1647057F6 (PROSTUT09), 2740029H1 (BRSTTUT14), 2740029X309F1 (BRSTTUT14), 2740029X321F1 (BRSTTUT14), 5068692H1 (PANCNOT23), 5193978H1 (LUNLTUT04), SBOAO3895D1
6	23	2414415	HNT3AZT01	1508631H1 (LUNGNOT14), 1596418F1 (BRAINOT14), 2414415F6 (HNT3AZT01), 2414415H1 (HNT3AZT01), 2414415X300D1 (HNT3AZT01), 2902794H1 (DRGCNOT01), 3080373H1 (BRAIUNT01), 3554689H1 (SYNONOT01), 3881949F6 (SPLNNOT11), 4996983H1 (MYEPTXT02)
7	24	2466714	THYRNNOT08	627910R6 (KIDNNOT05), 2358261R6 (LUNGTFET05), 2358464F6 (LUNGTFET05), 2466714H1 (THYRNNOT08), 2726053F6 (OVARTUT05), 3845383H1 (DENDNOT01), 4228854H1 (BRAMDIT01)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
8	25	2617942	GBLANOT01	1311045F6 (COLNFE02), 1375061F6 (LUNGNOT10), 1580318F6 (DUODNOT01), 2016443F6 (ENDCNOT03), 2617942F6 (GBLANOT01), 2703625H1 (OVARTUT10)
9	26	2945431	BRAITUT23	782666T6 (MYOMNOT01), 2715384T6 (THYRNNOT09), 2945431F6 (BRAITUT23), 2945431H1 (BRAITUT23), 4114919F6 (UTRSTUT07)
10	27	4074113	KIDNNNOT26	1754278T6 (LIVRTUT01), 1830341T6 (THP1AZT01), 2122328T6 (BRSTNOT07), 4074113F6 (KIDNNNOT26), 4074113H1 (KIDNNNOT26)
11	28	1413743	BRAINNOT12	1413743H1 (BRAINNOT12), 2643096F6 (LUNGUT08), SBWA04580V1, SBWA04306V1, SBWA00027V1, SBWA00447V1, SBWA00075V1, SBWA03280V1, SBWA00778V1
12	29	1733477	BRSTTUT08	1733477H1 (BRSTTUT08), 1733477F6 (BRSTTUT08), 3346594F6 (BRAITUT24), 1696051F6 (COLNNOT23), SAFC01922F1, 3234341F6 (COLNUCT03), 2634476T6 (COLNTUT15), SBCA03249F1
13	30	2641908	LUNGUT08	2641908H1 (LUNGUT08), SASA01215F1, SASB01010F1, SASB01616F1, SASA03523F1
14	31	2656554	LUNGUT09	2656554H1 (LUNGUT09), 2656554F6 (LUNGUT09), 2656554T6 (LUNGUT09)
15	32	2719228	THYRNNOT09	2719228H1 (THYRNNOT09), 2719228F6 (THYRNNOT09), 532553H1 (BRAINNOT03), 1626110F6 (COLNPOT01), 1626488F6 (COLNPOT01), 2100947R6 (BRAITUT02), ^a SBCA05174F1
16	33	3657824	ENDPNOT02	3657824H1 (ENDPNOT02), 2450345F6 (ENDANOT01), 2756551H1 (THP1AZS08), SCBA01425V1, 2394519T6 (THP1AZT01)
17	34	5378485	BRAXNOT01	5378485H1 (BRAXNOT01), SCAA06347V1, 1570276F1 (UTRSNOT05)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
1 384	S375 S34 T33 T89 T102 S161 S246 T298			E1-E2 ATPase signatures: R85-D115, T106-I128 P-type cation transporter signature: I109-I128 Transmembrane domains: I330-W351, L271-I288	E1-E2 ATPase [Mus musculus] (GI 2944187)	BLAST BLOCKS MOTIFS PFAM PRINTS
2 846	T80 S106 S22 T40 T109 S110 S141 S229 S245 S255 S288 T408 T614 T679 T707 S819 S24 S274 S293 T386 T450 T536 S568 T581 T610 T673	N301		ABC transporter signatures: F453-L467, G330-Q525, D652-Q816, L744-A758 ATP-binding (P-loop) domains: G337-T344, G659-S666	ABC transporter protein isoform [Homo sapiens] (GI 2522534)	BLAST BLOCKS MOTIFS PFAM PRINTS ProfileScan SPScan
3 511	S183 S99 S385 T387 T39 T106 S240 T376 T461 Y111 Y151 Y194 Y353	N195 N238 N258 N383		AMP-binding enzyme motif: R4-V404 AMP-binding domain signature: K91-M144 Lipocalin signature: E2-M14	Fatty acid transport protein [Homo sapiens] (GI 4206376) Fatty acid transport protein [Mus musculus] (GI 2612939)	BLAST MOTIFS PFAM PRINTS ProfileScan SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
4	718	S303 T224 T390 S501 T547 S560 S579 S623 T675 S248 S274 T326 S629 T655	N176 N229 N410	ABC transporter domain: L595-L609 ABC transporter family signatures: G483-V499, L595-D626 ABC transporter transmembrane domain: L128-M408 ABC transporter nucleotide binding domain: G483-G668 ATP/GTP-binding site motif A: G490-T497	ABC transporter protein [Homo sapiens] (GI 4321407) SMDR1 [Schistosoma mansonii] (GI 425474)	BLAST BLOCKS MOTIFS PFAM HMM
5	635	S25 T55 S128 T172 S242 S284 T378 T140 S222 S283 T286 T549	N138 N489 N498 N534	Transmembrane domain: F123-V144	Sodium:solute symporter family signature: M61-A467, N169-A216, T172-V217, G353-I383, I430-A467	BLAST BLOCKS PFAM ProfileScan HMM

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification / Homologies	Analytical Methods
6	535	T11 S61 S103 S179 S292 T305 S337 T363 S487 S502 Y109		Amino acid permeases signature: L36-K474 (from PFAM) Transmembrane domains: L76-L96, K111-Y131, S159-N176, I189-V207, N264-Y287, I422-F449, I451-W469	SLC7A8 [Homo sapiens] (GI 4581470) Amino acid transporter (GI 3639058, SEQ ID NO: 38)	BLAST BLOCKS MOTIFS PFAM
7	456	T106 S107 S215 S277 T306 T331 S419 S425		Aromatic amino acid permease signature: N53-S98 (from BLOCKS) Transmembrane domains: F47-V69, Y68-L88, T169-L191, L197-I216, S366-Y387, M330-I347, F432-I454	Aromatic amino acid transporter (GI 1840045, SEQ ID NO:39)	BLAST BLOCKS MOTIFS
8	325	T23 T29 S65 T84 S106 S133 S160		Transmembrane domains: I1209-D227, I1284-T307	UDP-N-acetylglucosamine transporter [Canis familiaris] (GI 3298605)	BLAST MOTIFS
9	178	S4 T22 S132 T151		ABC transporters signature: S2-G141 (from PFAM)	ABC transporter (GI 3335175, SEQ ID NO:41)	BLAST BLOCKS MOTIFS PFAM Profilescan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
10 255	T89 S107 S214			ABC transporters signature: Q45-G227 (from PFAM)	Anthracycline resistance-associated protein (ABC transporter; GI 1279457, SEQ ID NO:42)	BLAST BLOCKS MOTIFS PFAM Profilescan
11 462	T179 S11 S37 T114 S420 T174			Aromatic amino acid permease signatures: F58-A104, A243-F265, L375-K394	Transmembrane transporter [Schizosaccharomyces pombe] (GI 3367790)	BLAST MOTIFS PRINTS
12 758	T174 T251 S376 T615 S616 S633 S678 S713 S733 T138 S574 S603 T656 S687 T726 Y703		N167 N172 N614	Sulfate transporter signature: L1198-T508	Transporter [Mus musculus] (GI 5359730) Transporter [Homo sapiens] (GI 291964)	BLAST MOTIFS PFAM
13 336	T32 S99 S254 T274 S103 T104 T133 S273		N116	Transmembrane domains: L1211-G230, A175-L196, W42-F61, F139-F159	Organic anion transporter OATP-B [Homo sapiens] (GI 5006263)	BLAST HMM MOTIFS
14 103	S9 S92 Y14		N81	Transmembrane domain: V22-F41	Similar to Human Na+/H+ exchanger 2 (A57644) [Homo sapiens] (GI 1665827)	BLAST HMM MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification/ Homologies	Analytical Methods
15	123	S21 S118 T119	N39	Signal peptide: M1-S21	Similar to Sugar transporter [Caenorhabditis elegans] (GI 3878537)	BLAST HMM MOTIFS
16	222	T83 S61 S74			E. coli cation transport protein (GI495778)	BLAST MOTIFS
17	111		N100		Vacuolar proton-ATPase subunit M9.2 [Homo sapiens] (GI 2584789)	BLAST MOTIFS

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
18	Reproductive (0.292) Nervous (0.167) Hematopoietic/Immune (0.125)	Cell proliferative (0.750) Inflammation (0.208)	PSPORT1
19	Reproductive (0.264) Nervous (0.155) Hematopoietic/Immune (0.145)	Cell proliferative (0.691) Inflammation (0.236)	PINCY
20	Reproductive Nervous Gastrointestinal	Cell proliferation (0.66) Inflammation and immune response (0.24)	PINCY
21	Reproductive Nervous Gastrointestinal Hematopoietic/Immune	Cell proliferation (0.73) Inflammation and immune response (0.21)	PSPORT1
22	Nervous Reproductive	Cell proliferation (0.76) Inflammation and immune response (0.33)	PINCY
23	Reproductive (0.333) Nervous (0.296)	Cancer (0.519) Inflammation (0.185)	
24	Reproductive (0.250) Musculoskeletal (0.179) Hematopoietic/Immune (0.143)	Cancer (0.429) Inflammation (0.286)	
25	Reproductive (0.360) Gastrointestinal (0.200) Cardiovascular (0.120)	Cancer (0.520) Inflammation (0.160)	,
26	Nervous (0.267) Endocrine (0.133)	Cancer (0.533) Inflammation (0.333)	

Table 3 (cont.)

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
27	Urologic (0.333) Cardiovascular (0.167) Developmental (0.167)	Cancer (0.667) Inflammation (0.167)	
28	Nervous (0.222) Hematopoietic/Immune (0.185) Reproductive (0.185)	Cell Proliferation (0.481) Inflammation (0.407)	pINCY
29	Gastrointestinal (0.286) Reproductive (0.257) Nervous (0.200)	Cell Proliferation (0.743) Inflammation (0.286)	pINCY
30	Cardiovascular (0.400) Nervous (0.200) Reproductive (0.200) Gastrointestinal (0.200)	Cell Proliferation (0.800) Inflammation (0.200)	pINCY
31	Hematopoietic/Immune (0.500) Reproductive (0.250) Cardiovascular (0.250)	Cell Proliferation (0.750) Inflammation (0.500)	pINCY
32	Gastrointestinal (0.304) Hematopoietic/Immune (0.174) Nervous (0.174)	Cell Proliferation (0.826) Inflammation (0.347)	pINCY
33	Cardiovascular (0.238) Nervous (0.190) Hematopoietic/Immune (0.143)	Cell Proliferation (0.571) Inflammation (0.473)	
34	Nervous (0.510) Cardiovascular (0.102)	Cell Proliferation (0.612) Inflammation (0.266)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Comment
18	BRSTTUT03	Library was constructed using RNA isolated from breast tumor tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated multicentric invasive grade 4 lobular carcinoma. The mass was identified in the upper outer quadrant, and three separate nodules were found in the lower outer quadrant of the left breast. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular disease, coronary artery aneurysm, breast cancer, prostate cancer, atherosclerotic coronary artery disease, and type I diabetes.
19	LUNGUT12	Library was constructed using RNA isolated from tumorous lung tissue removed from a 70-year-old Caucasian female during a lung lobectomy of the left upper lobe. Pathology indicated grade 3 (of 4) adenocarcinoma and vascular invasion. Patient history included tobacco abuse, depressive disorder, anxiety state, and skin cancer. Family history included cerebrovascular disease, congestive heart failure, colon cancer, depressive disorder, and primary liver cancer.
20	BLADNOT06	Library was constructed using RNA isolated from posterior wall bladder tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder and urothelium. Patient history included lung neoplasm. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
21	PROSNON01	Normalized prostate library was constructed from 4.4 million independent clones from a prostate library. Starting RNA was made from prostate tissue removed from a 28-year-old Caucasian male. The normalization and hybridization conditions were adapted from Soares, M.B. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
22	BRSTTUT14	Library was constructed using RNA isolated from breast tumor tissue removed from a 62-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated an invasive grade 3 (of 4), nuclear grade 3 (of 3) adenocarcinoma, ductal type. Ductal carcinoma in situ, comedo type, comprised 60% of the tumor mass. Metastatic adenocarcinoma was identified in one (of 14) axillary lymph nodes with no perinodal extension. Tumor cells were strongly positive for estrogen receptors and weakly positive for progesterone receptors. Patient history included benign colon neoplasm, hyperlipidemia, cardiac dysrhythmia, and obesity. Family history included atherosclerotic coronary artery disease, myocardial infarction, colon cancer, ovarian cancer, lung cancer, and cerebrovascular disease.
23	HNT3AZT01	Library was constructed using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated for three days with 0.35 micromolar 5-aza-2'-deoxycytidine (AZ).
24	THYRNOT08	Library was constructed using RNA isolated from the diseased left thyroid tissue removed from a 13-year-old Caucasian female during a complete thyroidectomy. Pathology indicated lymphocytic thyroiditis.
25	GBLANOT01	Library was constructed using RNA isolated from diseased gallbladder tissue removed from a 53-year-old Caucasian female during a cholecystectomy. Pathology indicated mild chronic cholecystitis and cholelithiasis with approximately 150 mixed gallstones. Family history included benign hypertension.
26	BRAITUT23	Library was constructed using RNA isolated from left posterior brain tumor tissue removed from a 36-year-old male during a cerebral meninges lesion excision. Pathology indicated meningioma. Family history included malignant skin melanoma, atherosclerotic coronary artery disease, hyperlipidemia, Huntington's chorea, and rheumatoid arthritis.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
27	KIDNNOT26	Library was constructed using RNA isolated from left kidney medulla and cortex tissue removed from a 53-year-old Caucasian female during a nephroureterectomy. Patient history included hyperlipidemia, cardiac dysrhythmia, metrorrhagia, cerebrovascular disease, and atherosclerotic coronary artery disease.
28	BRAINOT12	Library was constructed using RNA isolated from brain tissue removed from the right frontal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), which are consistent with chronic seizure disorder. The patient presented with intractable convulsive epilepsy. Family history included a cervical neoplasm.
29	BRSTTUT08	Library was constructed using RNA isolated from breast tumor tissue removed from a 45-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive nuclear grade 2-3 adenocarcinoma (ductal type), with 3 of 23 lymph nodes positive for metastatic disease. Greater than 50% of the tumor volume was in-situ, both comedo and non-comedo types. There were also positive estrogen/progesterone receptors and uninvolved tissue showing proliferative changes. Patient history included valvuloplasty of mitral valve without replacement, rheumatic mitral insufficiency, and rheumatic heart disease. Family history included acute myocardial infarction and atherosclerotic coronary artery disease and type II diabetes.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
30	LUNGUT08	Library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fiberoptic bronchoscopy. Pathology indicated a grade 3 adenocarcinoma, forming a mass penetrating the pleura of the right upper lung lobe. The bronchial margin of the resection was free of tumor. The tissue from the superior segment of the right lower lobe lung showed multiple (2) calcified granulomas. Multiple lymph nodes were negative for tumor. In addition, budding yeast forms resembling histoplasma identified in two lymph nodes stained with GMS (silver). Patient history included atherosclerotic coronary artery disease, an acute myocardial infarction, rectal cancer, an asymptomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, lung cancer, type II diabetes, atherosclerotic coronary artery disease, and acute myocardial infarction.
31	LUNGUT09	Library was constructed RNA isolated from lung tumor tissue removed from a 68-year-old Caucasian male during segmental lung resection. Pathology indicated invasive grade 3 squamous cell carcinoma in the right upper lobe, forming an infiltrating mass involving the bronchus and the surrounding parenchyma. One (of 4) intrapulmonary peribronchial lymph nodes contained a metastatic tumor. An apical cap was identified. One (of 15) right paratracheal lower lymph nodes contained a metastatic tumor. Permanent superior mediastinal sections revealed metastatic squamous cell carcinoma in the lymph nodes. Patient history included of type II diabetes, thyroid disorder, depressive disorder, hyperlipidemia, esophageal ulcer, and tobacco use. Family history included brain cancer and atherosclerotic coronary artery disease.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
32	THYRNOT09	Library was constructed using RNA isolated from diseased thyroid tissue removed from an 18-year-old Caucasian female during an unilateral thyroid lobectomy and regional lymph node excision. Pathology indicated adenomatous goiter. This was associated with a follicular adenoma of the thyroid. The right neck lymph nodes displayed reactive follicular hyperplasia. The patient presented with hypophosphatemia. Patient history included compression of the brain, headache, iron deficiency anemia, active rickets, epidermal nevus syndrome (treated with lasers), and osteitis deformans. Family history included thyroid cancer, type II diabetes, benign hypertension, and nodular lymphoma.
33	ENDPNOT02	Library was constructed using RNA isolated from pulmonary artery endothelial cells removed from a 10-year-old Caucasian male. The cells were treated with TNF alpha and IL-1 beta 10ng/ml each for 20 hours.
34	BRAXNOT01	Library was constructed using RNA isolated from cerebellar tissue removed from a 70-year-old male. Patient history included chronic obstructive airways disease and left ventricular failure.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	Mismatch <50%
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	<i>ESTs:</i> Probability value= 1.0E-8 or less <i>Assembled ESTs:</i> fasta Identity= 95% or greater and Match length=200 bases or greater, fastx E value=1.0E-8 or less <i>Full Length sequences:</i> Probability value= 1.0E-10 or less
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	<i>ESTs:</i> fasta E value=1.0E-6 <i>Assembled ESTs:</i> fasta Identity= 95% or greater and Match length=200 bases or greater, fastx E value=1.0E-8 or less <i>Full Length sequences:</i> fastx score=100 or greater
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	<i>ESTs:</i> fasta E value=1.0E-6 <i>Assembled ESTs:</i> fasta Identity= 95% or greater and Match length=200 bases or greater, fastx E value=1.0E-8 or less <i>Full Length sequences:</i> fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger, and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 182:146-159; Barroch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Normalized quality score≥GCG-specified “HIGH” value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phil's Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater, Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. <u>supra</u> ; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	